22 OCT 1976

MEMORANDUM FOR: Deputy Director for Applications/ODP
Deputy Director for Processing/ODP
Chief, Management Staff/ODP
Chief, Special Projects Staff/ODP
Executive Officer/ODP

FROM : Clifford D. May, Jr.
Director of Data Processing

SUBJECT : Digital Processing of CDS Output for DDO

REFERENCES : a. My memo of 8 Oct 76 to Same Adses, Same Subj.
b. Memo to D/ODP fm C/ISG dtd 13 Sept 76, Subj: Cable Dissemination System Interface Planning
c. Memo to D/ODP fm C/ISG dtd 23 Aug 76, Subj: DDO CDS Traffic Requirements

1. Please disregard the action I requested in reference a. The ISG requirements described in reference c, which was forwarded to you earlier, and reference b, copy attached, will require the design of an ADP system which includes features similar to those in the SAFE system. I would like to have a brief discussion on the issues that this situation raises at 1430 hours on 4 November in 2D-03. In particular, I would like you to be prepared to discuss:

a. What commonality exists between the ISG and SAFE requirements for electrical processing of digital cable traffic from CDS?

b. Can the SAFE system design be modified to accommodate the ISG requirement? If so, what would be the impact on the SAFE requirements and on costs?

c. What justification would we have to pursue the development of the ISG system separately from SAFE? What are our preliminary estimates of time and cost for the ISG development?

2. At the conclusion of the meeting, we will discuss the follow-up action that should be taken on this matter.

[redacted] STAT

Clifford [redacted]ay, Jr. STAT

Att

Distribution:
One - ea Adse
2 - O/D/ODP

STAT O/D/ODP/[redacted]ksk (22 Oct 76)

ROUTING AND RECORD SHEET

SUBJECT: (Optional)
Cable Dissemination System Interface Planning.

ODP # 1838-76

25X1

FROM: DDO/C/SS/
1D 23

EXTENSION

NO. 25X1

DATE

TO: (Officer designation, room number, and building)	DATE RECEIVED	DATE FORWARDED	OFFICER'S INITIALS	COMMENTS (Number each comment to show from whom to whom. Draw a line across column after each comment.)
1. Director of Data Processing 2D 00 DX-6 Green				
2. ED		9/14	CD	Copy to DD/A/ODP Attn D Division Done 9/16. CD CD
3.				
4.				
5.				
6.				
7.				
8.				
9.				
10.				
11.				
12.				
13.				
14.				
15.				

ODP #1838-76

13 SEP 1976

MEMORANDUM FOR: Director of Data Processing

FROM : [redacted] 25X1
Services Staff, DDO

SUBJECT : Cable Dissemination System Interface Planning.

25X1

1. At this morning's CAMS briefing (which was excellent, by the way, and greatly appreciated) we mentioned the significance which we attach to the collection of messages in digital form, drawing on the output of the Cable Dissemination System. Attached is a copy of a recent report prepared by [redacted] on the planning done to date within [redacted] for the exploitation of CDS. As you will see it is related to my memorandum of 19 August to you and the Director of Communications, subject: DDO CDS Traffic Requirements. 25X1 25X1

25X1

2. If you or any of your staff have any questions regarding this paper, feel free to call [redacted]. We will be happy to meet with you any time to discuss these or any other ADP matters of common concern.

25X1

[redacted]

Attachment: a/s

[redacted] 25X1

IS/AD-76/101
18 August 1976

25X1

MEMORANDUM FOR: [redacted] 25X1

FROM : [redacted]
Software Specialist,
[redacted] 25X1

SUBJECT : Cable Dissemination System
Interface Planning

Attached is a report which describes current [redacted] 25X1
planning pertaining to the proposed DDO interface with
the Cable Dissemination System. The report documents
the briefing presented to you on 12 July. A copy has
25X1 been forwarded to [redacted] as you requested.

25X1

[redacted]

Attachment: A/S

[redacted] 25X1

CABLE DISSEMINATION SYSTEM INTERFACE PLANNING

CABLE DISSEMINATION SYSTEM INTERFACE PLANNING

INTRODUCTION

This report documents the results of a study by [redacted] STAT which examined the concept of a DDO digital interface with the impending Cable Dissemination System (CDS). The overall purpose of the study was to analyze CDS capabilities in terms of their potential application in the DDO central records system. In particular, the analysts attempted to focus on the following areas:

°What is the CDS?

- -What will be the CDS-DDO interface?
- -What type of traffic will come through CDS to the DDO -- how will it come -- how much will come -- how will it be distributed over time?

°What is the optimum use the DDO records system can/should make of a CDS interface?

- -Could DDO destined CDS traffic be received and stored in digital form?
- -If so, would it be feasible to store, process and retrieve all such traffic digitally, or would storage space requirements preclude this?

°How would a digital storage capability fit into the current automated DDO records system?

- -What would be its relationship to STAR -- to WALNUT -- to the Document Reference System (DRS) -- to the users?

°How would a digital storage capability fit into a future automated DDO records system?

- -What would be the long range effects on ADP plans -- on the user -- on the modus operandi -- on operational capability?

what is its relationship to SAFE? (Ray)

°How should such a capability be phased into operation?

-How would it affect existing planning?
-Can it be broken up into manageable portions?
-What are the magnitudes of the efforts involved, particularly in terms of cost, schedule, manpower, and risk?

The report is presented in four parts. The first part, entitled Background, attempts to "set the stage" for the following analysis of the CDS interface. Here, the authors philosophize about the advent of CDS and provide a commentary on its potential importance and impact on operational capability.

The second section, called The Big Picture, establishes the context for the CDS analysis by addressing how the interface would fit into the overall scheme of things; it attempts to place the subject in its proper perspective. Here the authors postulate a set of long term ADP goals related to the advent of CDS, and thus show how the near term CDS objectives lead naturally to a logical and purposeful end -- the development of a future system which reflects not only long range system growth in parallel with advancing technology, but also a continual improvement in operational capability consistent with the needs and desires of the user community.

The third part of the report, Near Term Objectives, focuses on those opportunities amenable to immediate accomplishment. This section provides a further explanation of these objectives, and sketches a proposed methodology for functionally relating them with the existing STAR system.

In the fourth section, Planning Estimates, the study addresses the size of the effort described in part three, based on an arbitrary time period of 10 years. The development activities are broken down into their basic elements, and each element is sized in terms of man months, calendar months, and estimated completion date. In this section also, the report presents an estimate of the CDS traffic input rate, and analyzes it in terms of the amount of auxiliary storage required to store it on-line for a period of 10 years.

Early in the study, it was recognized that no project of this magnitude and import should exist without a name and the requisite catchy acronym! Accordingly, the selection of a suitable name became a high priority endeavor.

Consistent with the celestial aura attendant to the (STAR) system with which this project must interface most closely, the name COMET leaped to the fore. COMET, of course, stands for the Collection of Operational Messages Electrically Transmitted.

It is important to the understanding of this report that the reader know a few fundamental facts about CDS, a message distribution system operated for the Headquarters by the Office of Communications. Its primary purpose is to improve the dissemination process of documents received in, and transmitted from, the Headquarters by electrical means. To the extent possible, this improvement will be achieved through automation. In any event, the dissemination will be determined, the documents will be addressed accordingly, and then distributed to the assigned addressees in hard copy form. This process, although an important one, does not represent a new interface to the DDO records system; only the dissemination process is improved. The key factor in the CDS process is its ability to also disseminate traffic in digital form. Thus, in addition to distributing the documents in hard copy form, CDS will also pass to ODP a requested subset of the traffic in digital form. While the majority of this traffic is expected to be destined for CRS, it can include whatever DDO traffic is requested. ODP will then extract the DDO subset from the CDS digital stream and pass it to [redacted] initially via magnetic tape. This study is concerned with the DDO portion of the CDS digital stream received via ODP. It should be noted that the DDO will have exclusive use of this subset of the CDS traffic.

STAT

BACKGROUND

In the initial phase of the study the analysts attempted to view the potential CDS interface as broadly as possible in an effort to gain an overall perspective concerning what this new capability would mean to the DDO central records system. It was soon recognized that CDS indeed represented a very profound change. For the first time, records to be maintained in the system could be received by the DDO in digital form. This was quickly perceived to be the key finding concerning the potential value the CDS interface would have to the records system. Indeed, it was quite obvious that the past inability of the DDO to receive the documents in digital form had had a dramatic influence on the design of the ADP system supporting the records system.

Primarily because it was not feasible to process the documents themselves, the system was designed to process information within and about the documents, such as biographic data concerning persons mentioned in the text (index records), bibliographic information about the documents (abstract records), and document locator information (DRS input). The documents themselves, however, were handled outside the ADP system, either in the filmed store (WALNUT) or in hard copy form. With the advent of the CDS interface, however, it would be possible for the first time to process the documents themselves, not just information about the documents.

THE BIG PICTURE

The analysts recognized that the CDS interface should not be viewed in isolation; rather, it should be examined in the context of other related systems, activities, plans, and objectives. It was clear that the CDS interface was only part of a larger whole, and that its development should be supportive of, and integrated with, the larger entity. Thus, an effort was made to gain an insight into the overall scheme of things in which the CDS interface would fit. The methodology chosen was to identify the set of near term and long term goals and objectives which could reasonably be related to the advent of CDS. In this way, the CDS interface could be viewed as a means to an end. Also, the associated overall development effort could be logically broken into phases which were not only manageable in size, but which also represented reasonable incremental improvements in functional capability. The overall results of this analysis are highlighted in Figure 1.

In the preliminary stages of this analysis, the study team addressed the question of the proper relationship of COMET with ALLSTAR. It was felt that although COMET should be functionally integrated with ALLSTAR, it should remain a separate entity to the extent possible. This would preclude the growth of ALLSTAR to unmanageable proportions, and would enhance the reliability and maintainability of both the COMET and ALLSTAR systems.

The first phase includes those near term objectives worthy of immediate attention, and is forecast for completion during 1978. Phase II represents what is perceived to be the

THE BIG PICTURE

°Phase I (1978)

-Receive and store digital stream on tape

--When CDS becomes operational

--Not yet integrated into ALLSTAR

--To capture information for future use

-Abstract, index and file in digital store

--Now functionally integrated into ALLSTAR

-Retrieve and print from digital store

--First operational use of digital store

°Phase II (1980)

-On-line analysis

--By IP analysts

°Future Phases (post 1980)

-On-line document retrieval

--By end users

--From digital store

--From non-digital store

-On-line document preparation and transmission

-On-line information retrieval

FIGURE 1. THE BIG PICTURE

next logical step in the development effort, and is planned for completion in 1980. Additional objectives are not phased in any definitive way, but are merely articulated for the purpose of understanding where COMET will likely lead. These "future phases" can be expected to be implemented in the post-1980 time frame.

PHASE I

Only a cursory analysis was necessary to recognize that the most immediate objective of the COMET project should be to receive and store the digital stream of traffic to be routed to the DDO via ODP. This action would preclude the loss of this digital information into the "bit bucket"; a needless waste. This digital stream is scheduled to begin arriving in December 1976 or January 1977, only a few months hence. Although there is not sufficient time to develop the software and procedures necessary to integrate this traffic into the ALLSTAR system by that time, we can at least be prepared to receive and store it for potential future use. There are several applications in which this data could be very useful. One example would be to use it for document storage and retrieval, after Phase I became operational. This could be done by transferring the digitized documents from the tapes to the ALLSTAR digital store, updating DRS with their location information, and "purging" their filmed images from the WALNUT system. Another example would be to use these digital records in a future conversion of the WALNUT system. It would be considerably easier to convert from digital form than from microform, particularly if the conversion was to a digital system.

Whereas, the initial objective should be simply to capture and save the digital stream, the next objectives are clearly the development of the facilities necessary to integrate the processing of digitized documents into the ALLSTAR system. That is, to be able to abstract and index digitized records, file them in a digital store, then retrieve them from the digital store and print them out upon demand.

PHASE II

Given that the digitized messages were integrated into the ALLSTAR system, an objective which could be expected to be realized by 1978, attention should then be focused on the concept of putting the records system users on-line to the computer. This concept, known variously as on-line processing,

remote processing, etc., basically reflects the idea of moving the ADP support from the computer room out to the user at his/her location. The first such activity proposed is that of the analysis function; the preparation of the abstract and index records by the [redacted] analysts. This then becomes the next objective, and constitutes what is referred to as Phase II. It is forecast for completion in 1980. The concept is to locate keyboard-CRT type terminals at the [redacted] analysts' locations and provide them computer assistance in their message analysis functions. This assistance could be provided in a number of ways, as indicated below:

- Provide the capability to view the incoming documents on the CRT screen and indicate the abstracting/indexing desired through the use of such facilities as function keys, light pen, cursor, etc. In other words, provide for on-line annotation of the displayed documents. Or,

- Provide a split screen capability with the incoming message displayed on half the screen, and the other half available to key in abstract/index records. In other words, permit the [redacted] analysts to compose the actual index/abstract records at the terminal rather than merely provide annotations for later record composition.

- Provide the capability for the analyst to enter a 201 file number and retrieve the associated true name, and vice versa.

- Provide the capability for the analyst to enter a file name or number and retrieve a listing of the file contents.

- Provide the capability for the analyst to enter a document identifier or a name and retrieve the associated index and abstract records. This would permit the analysts to review the indexing/abstracting process of documents associated with the one in process.

Note: The last three capabilities already exist in ALLSTAR for other users, so it would only be necessary to extend them out to the [redacted] analysts at their locations.

FUTURE PHASES

The additional ADP goals related to the CDS development -- on-line document retrieval, on-line document preparation and transmission, and on-line information retrieval -- are also clearly supportive of the previously mentioned philosophy of putting the DDO users on-line.

On-line document retrieval is an objective which brings the records system end-user on-line. Here, the area division and staff personnel who normally receive hard copy documents as a result of name trace activity are provided direct computer support for the first time. That support would be in the form of a terminal with which they could directly order and retrieve document images. These document images could be in one or more of several forms, as indicated below:

- °CRT display of messages received electrically via CDS and stored digitally; perhaps with a screen-to-printer option.
- °Hard copy printout of messages received electrically via CDS and stored digitally.
- °Video (raster scan) display of document images stored in non-digital form (probably microform). These would probably be documents which were either input to the records system prior to the CDS implementation, or were input in hard copy form.
- °Hard copy facsimile printout of document images stored in non-digital form.

In addition to providing document images to area division and staff personnel, the computer terminals could also be used to provide additional information, such as:

- °Locator information on hard copy and filmed documents.
- °Restriction information on sensitive documents, e.g. who or what office controls access.

It is obviously much easier to provide for the remote display of digitized documents than non-digitized documents. This objective could thus be expected to be met incrementally, with the initial capability providing for the remote display of only digitized documents, and the final capability providing

for the remote display of non-digitized documents. There may be non-technological considerations also, such as whether all document recipients really need or desire direct document retrieval.

The next objective proposed is a natural extension of the previous one. Given that the end-user, or so called "division analyst", has a keyboard-CRT type terminal at his/her location, it would seem logical to provide for the use of that terminal to compose and authenticate messages and to release and transmit them via CDS. In other words, once connectivity is established with the end-user, provide the capability to use CDS in the reverse direction. It is not anticipated that much, if any, technological development need be associated with this objective; once the user is provided a terminal for document retrieval, that terminal and its associated communications capability may facilitate document preparation and transmission with little or no upgrade.

On-line information retrieval represents the ultimate CDS related ADP objective which was envisioned in this study. The word information is key here, and the capability for information retrieval is contrasted with that of document retrieval. Obviously, information retrieval represents a much more broad and powerful capability than document retrieval. It reflects a synthesizing of the contents of the documents into a data base of information supported by a highly sophisticated user interface and retrieval system. The point is, though, that the computer system would now be used to provide analytical support to the user in addition to the retrieval of stored records. Significantly, it is CDS which first makes this powerful capability possible. Since it provides the documents in digital form, the documents may themselves be analyzed and processed by the computer system. Their contents can be extracted and manipulated as required to construct a data base which can support an information retrieval system. On-line information retrieval will be difficult and complex and will be a long time in coming, but it should certainly be recognized as a long range goal and may represent the most significant benefit of CDS to the DDO.

NEAR TERM OBJECTIVES

Having set the long range course for COMET development, the analysts turned their attention to those near term objectives which require immediate attention. These objectives,

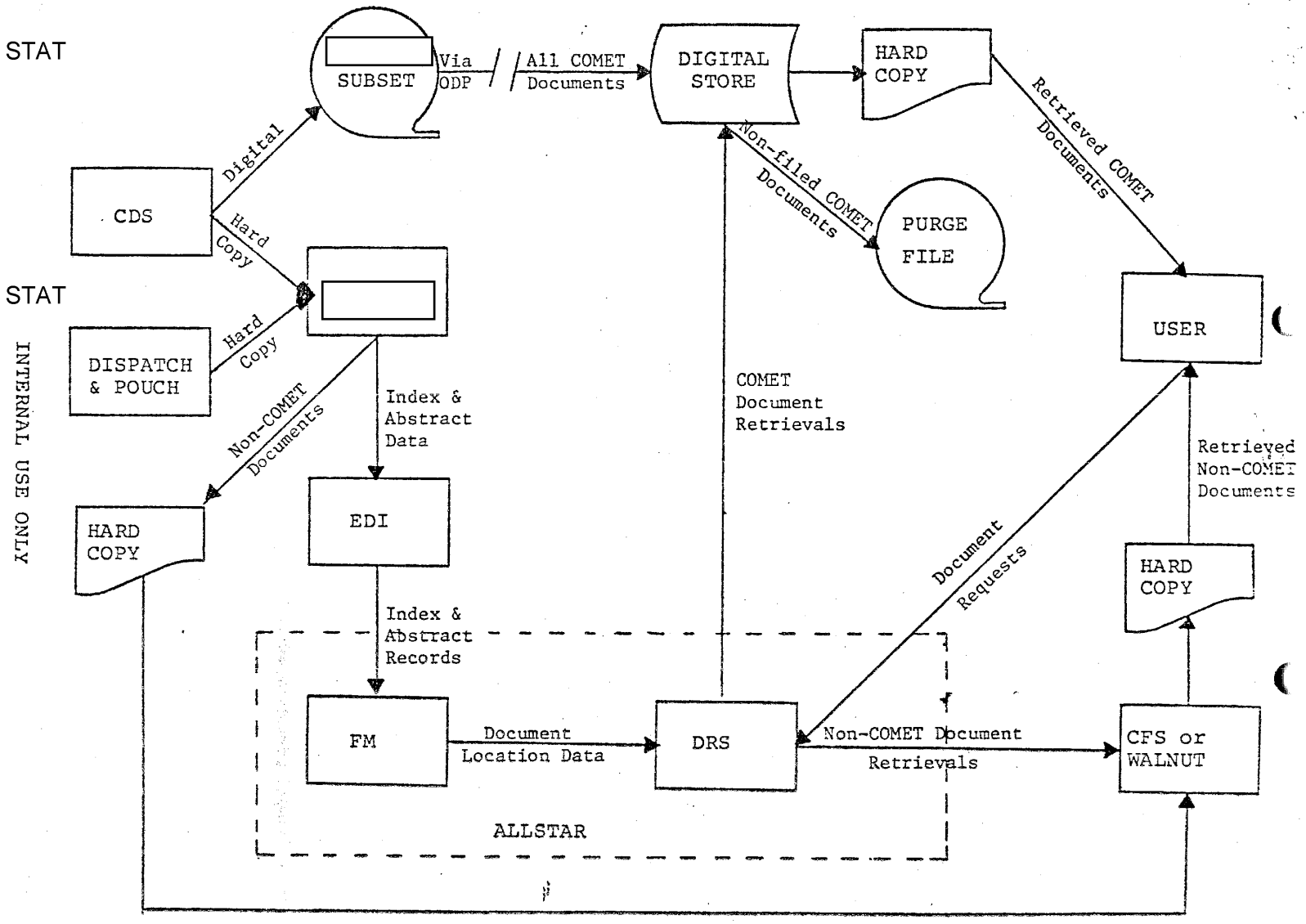
comprising what is known as Phase I, are first simply to receive and store the DDO digital stream received from CDS via ODP, then to integrate the incoming digitized documents into the ALLSTAR system by indexing/abstracting them and filing them in a digital store with the capability to retrieve and print them upon request.

The first near term goal, to receive and store the digital stream from the outset represents a relatively simple, straightforward task. It does not connote any immediate operational use of this early digital traffic, since it does not include any integration of the initial CDS traffic into the ALLSTAR system. It is only an interim step, and basically reflects an effort to capture and retain the incoming "bits" while the operational COMET system is being developed (although this data may later be integrated into ALLSTAR). During this time, the non-digitized (hard copy) CDS traffic will be processed into the ALLSTAR system just as the pre-CDS hard copy traffic was. Meanwhile, CDS will pass the digital form of the DDO traffic to ODP via its (one and only) 9600 baud data link to the general computer center in room GC-03. The DDO traffic will be only a part of the larger digital stream carrying Agency-wide traffic on the single data link. Therefore, ODP must identify and extract the DDO portion before writing it onto magnetic tape. Periodically, the tapes will be transferred to the special computer center in room GC-47, and saved for whatever future use is deemed appropriate. An important fact concerning this first COMET capability is that it must be available when CDS first becomes operational. However, as it is a rather trivial task, there would seem to be little risk associated with achieving such an objective.

The remaining near term goal is to integrate the digitally received CDS traffic into the ALLSTAR system. That is, to abstract and index it, file it in a digital store (if it is to be saved), and provide a capability to retrieve and print it upon request. Although the study team has not prepared a detailed design showing how such a capability would be implemented, it has postulated an overall system flow (see Figure 2.)

The first part of the flow can be viewed as the input end; that is, the receipt, abstracting and indexing, and filing of the incoming documents. To begin, CDS would output the DDO electrical traffic in two forms: hard copy and digital. The hard copy form would be routed to [redacted] for analysis just as it would if no digital form existed (and just as it is now). The digital form, meanwhile, would be routed from CDS via data

STAT



link to the GC-03 computer center then to the GC-47 computer center via magnetic tape and placed in digital store. The IP analysts would examine the hard copy version of the documents and determine whether they should be kept. If so, the analysts would annotate the documents with the appropriate inscriptions for later composition of abstract and (where necessary) index records by Electronic Data Input (EDI) personnel. If the analysts determined it was not necessary to keep (file) the documents, they would be discarded (additional copies would have already been distributed to the other DDO addressees). The documents selected for retention and filing, however, would be passed to EDI personnel who would compose and key-in abstract and index records from the annotations made by the [redacted] analysts. These records would then be entered into the ALLSTAR system via the nightly file maintenance (FM) computer run. During the FM processing, the ALLSTAR Document Reference Subsystem (DRS) would be updated to add these document identifiers to its inventory list, and to indicate that they were located in the digital store. Since all CDS documents received by the DDO would be stored in the digital store, and only a subset selected for retention by the IP analysts, the digital file would contain extraneous document images. Thus, special processing would frequently be required to compare the inventory of documents actually filed (the DRS inventory) with the inventory in digital store. Those extraneous documents in digital store would be purged to a special purge file kept indefinitely on magnetic tape. An extended period of time (perhaps six months) should be allowed to pass before documents would be purged, to permit ample opportunity for reversal in the initial decision not to retain the documents.

STAT

The second part of the system flow can be viewed as the output end; that is, the retrieval and printout of the requested documents. As users determined a need for the CDS documents, perhaps as the result of a name trace, they would submit a document request to the ALLSTAR system. DRS would locate the document identifier in its inventory list and recognize that it was located in digital store. DRS would then activate a retrieval program to locate and retrieve the digitized document from the digital store. The retrieved record would be formatted as necessary and output to the printer. The printed document would then be routed manually to the user. It is important to recognize that this "output end " would be developed in parallel with the aforementioned "input end". Significantly, when both development efforts

were completed, WALNUT filming could be stopped for all documents transmitted and received electrically.

One important consideration in planning for the CDS interface is the selection of the particular subset of CDS traffic which is to be routed to the DDO in digital form. In accordance with the system flow described above, the answer is clear -- the hard copy subset which is routed to IP/AN. Whatever criteria is used to establish that a certain subset of the CDS traffic should be routed to IP/AN in hard copy should also be used to select the subset to be routed to the DDO through ODP in digital form. In other words, the DDO digital stream should duplicate exactly the CDS hard copy traffic routed to IP/AN. Currently, this traffic could be characterized as all DDO operational electrical traffic except that which is top secret or prescribed and limited (P&L). One addition to this are those electrical Intelligence Information Reports which are indexed.

PLANNING ESTIMATES

In an attempt to quantify the development effort associated with each of the three parts of Phase I, the study team outlined a master system development plan identifying the specific work activities required and placing them in the proper sequence. This plan was then applied to each of the three Phase I activities to elicit schedule and manpower estimates attendant to the development efforts. The master plan is outlined below:

°System analysis

- -Requirements/feasibility analysis
- -Plan development
- -Plan approval
- -Staffing

°System design

- -Specifications preparation
- -Procedures development
- -Procedures coordination

°System development

-Software and files
-Hardware
-Documentation

°System integration and test

-Test plan preparation
-Test and debug activities
-Acceptance testing
-Parallel operation

The system analysis effort is basically devoted to determining the precise project objectives, estimating the work effort, and getting management approval. The initial analysis effort should be directed at defining the user requirements, ensuring they meet the test of feasibility. A plan should then be developed, mapping the achievement of incremental objectives against time, and indicating the resource requirements associated with each increment. This plan should then be submitted to management for approval. Once the plan is approved, the necessary technical and management personnel can be assigned to the project as their particular skills are required.

The system design phase is an effort to document the planned activities in sufficient detail to ensure that the system subsequently developed will indeed meet the user requirements identified in the system analysis phase. The principal element of this phase is the production of programming specifications. Other activities would be oriented toward developing the new operating procedures attendant to implementation of the new system and coordinating them with all persons involved.

The system development phase would normally represent the majority of the work effort, as it requires the preparation of new software and the modification of existing software. Other activities include the acquisition and installation of any new hardware required, and the preparation of programming documentation.

The final phase, system integration and test, reflects the test and debug activity associated with bringing all the development pieces together and making them function smoothly as a system. This phase is highlighted by an acceptance test

in which the users should participate to ensure that the finished product is sufficiently responsive to the requirements they articulated at the beginning of the project.

Figure 3 reveals the mapping of the first Phase I objective (receive and store) against the system development plan. It indicates a requirement for about two persons full time to ensure the completion of the necessary system development activities prior to the estimated CDS operational date of January 1977.

Figure 4 addresses the other two Phase I objectives. It reflects a 22 calendar month effort with the two objectives being pursued concurrently. It further shows an estimate of 57 man months to develop the capability to abstract and file the documents, and 51 man months to develop the retrieve and print capability. The largest single effort is estimated to be, as expected, the software development activity. It is estimated that three programmers would be required for 6 months on the first effort, and two for 6 months on the second.

To determine the hardware resources necessary to support the COMET project an estimate of the CDS traffic rate was obtained (see Attachment 1). This estimate indicates that about 540 million characters of the CDS traffic could be expected to be routed to IP/AN annually, and that about 400 million characters would be in the subset kept for abstracting and filing. Using this figure as a planning factor, a 10 year projection was made (Figure 5) to indicate how much disk storage would be required to accommodate all the CDS traffic to be retained by the DDO. The projection indicates that two IBM 3350 disk units (16 drives) would be required in 10 years, costing about [redacted] and requiring about 80 square feet of floor space. The projection, as reflected in column 3 of Figure 5, includes an adjustment to the 400 million characters per year of 10% for device formatting and 20% for indexing (column 3 equals 1.3 times column 2). The projection also reflects a 6 month delay before purging (column 2 carries 70 million additional characters per year throughout the entire period).

STAT

The total Phase I planning estimates are indicated below:

°Man months - 121

°Calendar months - 22

	RECEIVE & STORE		
	Man Months	Calendar Months	Completion Date
SYSTEM ANALYSIS			
°Requirements	1	1	7/76
°Planning	1	1	8/76
°Approval	-	-	8/76
°Staffing	-	-	8/76
SYSTEM DESIGN			
°Specifications	2	1	9/76
°Procedures	2	1	10/76
°Coordination	2	1	11/76
SYSTEM DEVELOPMENT			
°Software & Files	2	1	12/76
°Hardware	-	-	-
°Documentation	1	1	12/76
SYSTEM INTEGRATION/TEST			
°Test & Debug	2	1	1/77
TOTALS	13	7	1/77

FIGURE 3. PLANNING ESTIMATES

	ABSTRACT & FILE			RETRIEVE & PRINT		
	Man Months	Calendar Months	Completion Date	Man Months	Calendar Months	Completion Date
SYSTEM ANALYSIS						
°Requirements	2	2	8/76	2	2	8/76
°Planning	2	2	8/76	2	2	8/76
°Approval	-	3	11/76	-	3	11/76
°Staffing	-	-	11/76	-	-	11/76
SYSTEM DESIGN						
°Specifications	4	2	1/77	4	2	1/77
°Procedures	6	2	3/77	6	2	3/77
°Coordination	6	2	5/77	6	2	5/77
SYSTEM DEVELOPMENT						
°Software & Files	18	6	11/77	12	6	11/77
°Hardware	-	6	5/77	-	6	5/77
°Documentation	4	2	11/77	4	2	11/77
SYSTEM INTEGRATION/TEST						
°Test & Debug	15	5	4/78	15	5	4/78
TOTAL	57	22	4/78	51	22	4/78

FIGURE 4. PLANNING ESTIMATES

STORAGE ESTIMATES

Years	No. Bytes To Be Stored ($x10^9$)	No. Bytes Storage Required ($x10^9$)	No. 3350 Drives Required (8/unit)	Est. Monthly Rental ($x10^3$)	Floor Space Required (FT^2)
1	.47	.61	2		10.5
2	.87	1.13	4		20.3
3	1.27	1.65			
4	1.67	2.17	6		30.0
5	2.07	2.69	8		39.8
6	2.47	3.21	10		50.3
7	2.87	3.73			
8	3.27	4.25	12		60.1
9	3.67	4.77	14		69.8
10	4.07	5.29	16		79.6

FIGURE 5. STORAGE ESTIMATES

°Completion dates

-Receive and store - January 1977
-Abstract and file - April 1978
-Retrieve and print - April 1978

°Cost

-Hardware (disk) only

STAT

-- [redacted] (first year)
-- [redacted] (in 10 years)

CONCLUSION

In conclusion, it can be seen that the impending DDO digital interface with CDS represents a significant upgrade in the potential for providing automated support to the central records system. Since CDS will make available the DDO documents in digital form, the documents themselves can now be processed by the computer for the first time. This important development opens up powerful new vistas of improved computer support to records system users. The most significant near term improvement is the capability to store the digitized documents in the computer system, making them readily available for rapid retrieval and printout upon request. This study indicates that such a capability could be provided within two years at a cost of about 10 man years of effort. The study further shows that all DDO electrical traffic selected for retention could be stored in this way for at least 10 years with no difficulty. Further, continuing advancements in computer storage technology should not only permit a significant extension to that period but also an improved cost per bit ratio.

Additional improvements resulting from the capability to computerize the incoming documents focus on the concept of placing the records system users on-line to the computer. The study addresses this concept in terms of on-line document analysis, on-line document retrieval, and on-line document preparation and transmission. Finally, the study recognizes the ultimate benefit to be gained from processing the documents in the computer; that of synthesizing the contents of the various documents into a data base designed to support on-line information retrieval. The study recommends that such a capability be recognized as a long range goal, and that all interim development efforts be supportive of that concept.

ATTACHMENT I

TO CABLE DISSEMINATION SYSTEM INTERFACE PLANNING

REPORT

12 July 1976

25X1 MEMORANDUM FOR: [redacted]

VIA : [redacted] 25X1

25X1 FROM : [redacted]

[redacted] 25X1

SUBJECT : Estimates on CY 1976 Electrical Traffic for CDS Planning

1. In the course of our recent joint planning efforts, we discussed the necessity of obtaining an estimate on the amount of electrical traffic [redacted] would receive as a result of the expected inauguration of the Cable Dissemination System later this year. In view of this fact, and that [redacted] has been engaged in several studies involving statistics on staff traffic, I have assembled a set of figures which will confirm and enlarge on previous verbal estimates on initial storage requirements. 25X1 25X1

2. These figures were derived from projected traffic volumes for CY 1976 based on the average monthly electrical traffic for January through June 1976. They encompass all operational and intelligence cables and telepouch to and from the DDO subdivided into the following categories:

a. traffic routed to [redacted] for processing; 25X1

b. traffic selected by [redacted] for retention in the DDO Records System; 25X1

c. retained traffic classified as indexed (IN), abstracted and cross referenced (AC) or abstracted only (AO).

3. The following table constitutes a summary of the amount of electrical traffic, in millions of characters, for each of the categories mentioned above.

25X1

E2 IMPDET [redacted]

Category	Cable Characters	Telepouch Characters	Total Characters
DDO operational and intelligence staff traffic	626M	81M	707M
Traffic routed [redacted]	465M	74M	539M
Traffic selected for retention in the system by [redacted] *	349M	44M	393M
Retained traffic classified as (IN)	41M	5M	46M
Retained traffic classified as (AC)	101M	9M	110M
Retained traffic classified as (AO)	207M	30M	237M

25X1

25X1

* Excludes P&L and Top Secret, etc., but not RYBAT traffic which constitutes roughly 20% of the cables processed by [redacted] Also excluded is any non-Agency electrical traffic which is felt to be nominal.

25X1

** Excludes non-Agency electrical traffic.

4. Having established estimates for electrical storage, it is interesting to contrast the storage estimates for IN, AC and AO documents from the table above with recent statistics on document input and retrieval activity. The results, although not totally substantiated at this point, reveal interesting if not surprising facts concerning both future demands on the digital store and its organizational structure.

5. In this regard, several comparisons are possible. The first compares storage, input and retrieval for each type of document by allowing all documents regardless of age (origination date), while the second considers documents grouped by date of origination. For our purposes, it is sufficient to consider two intervals, (1970-1975) and (1965-1975). A summary of the results is presented in the tables below.

6. Although the figures on document input and retained document store can be substantiated by several recent studies, those concerning total document requests factored by age of

document are more elusive and had to be determined indirectly. Details on calculations and sources are available if you wish to see them.

Storage vs. Document Input and Retrieval Activity

a. All documents (1940-75):

	IN	AC	AO
Percent of total document input	25	10	65
Percent of total documents requested	70-75	10-15	15
Percent of retained electrical store	12	28	60

b. Documents originated (1970-1975):

Percent of total document input	7-10	16-19	71-77
Percent of total documents requested	22	3	5
Percent of retained electrical store	12	28	60

c. Documents Originated (1965-1975)

Percent of total document input	7-10	16-19	71-77
Percent of total documents requested	38	5	7
Percent of retained electrical store	12	28	60

7. Percentages in table (a) relate to properties of a document storage system (such as our present combined storage system of hardcopy, and microforms) capable of satisfying 100% of retrieval demands. The percentages in table (b) theoretically exemplifies the properties of an electrical storage system six years after implementation in 1970. Table (c) characterizes a similar electrical store 11 years after implementation. Note that the electrical store started in 1970 is capable of satisfying at most 30% of the total request load in 1975. This figure improves to 50% for the electrical store inaugurated in 1965. In accordance with present plans, these results seem to indicate that implementation of on-line document retrieval from digital store would be premature much before 1986. I might mention that the estimates on the percentage of retrieval from electrical stores may prove slightly optimistic since only electrical traffic (cables and telepouch) could be retrieved under the present constraints on input. OCR or NCI conversion of the retrospective hardcopy or microfilm store or future expansion on CDS traffic routed to [redacted] could, of course, drastically

25X1

alter retrieval percentages. Unfortunately, for electrical stores, unlike microstores, it is far more difficult to reach back into the past to gather material for initial loading. The payoff for the waiting is of course greater organizational flexibility and ease of update and purge. For these reasons it appears that the Walnut data base, or a portion thereof, will of necessity remain with us for a number of years since retrieval of this material in hardcopy form from IP/CFS would impose a considerable added burden, especially in view of probable future resource constraints.

8. As to the organizational implications, the numbers speak for themselves. Indexed material organized by date of origination is probably the most likely initial candidate for grouping.

9. As additional data becomes available, I expect to be able to confirm or refute figures which were derived from only partially substantiated assumptions.

[redacted]

25X1

25X1

Distribution:

Original - [redacted]
1 - [redacted]
1 - [redacted]

25X1

25X1

25X1